BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 3: e37221
doi: 10.3897/biss.3.37221

OPEN ACCESS

Conference Abstract

The Catalogue of Life: Assembling data into a global taxonomic checklist

Geoffrey Ower‡, Yuri Roskov‡

‡ Illinois Natural History Survey, University of Illinois at Urbana-Champaign, Champaign, IL, United States of America

Corresponding author: Geoffrey Ower (gdower@illinois.edu)

Received: 12 Jun 2019 | Published: 19 Jun 2019

Citation: Ower G, Roskov Y (2019) The Catalogue of Life: Assembling data into a global taxonomic checklist. Biodiversity Information Science and Standards 3: e37221. https://doi.org/10.3897/biss.3.37221

Abstract

Producing a global taxonomic checklist of all species is essential for indexing biodiversity data, and for providing the basic knowledge needed to study, manage, and conserve biological diversity. The Catalogue of Life (CoL) aims to provide a global taxonomic checklist of all species, and includes 1.9 million species names in the 2019 annual edition. The task of assembling data into CoL is complex and requires reformatting data, quality assurance testing, and collaborating with data providers to resolve detected taxonomic conflicts. Global Species Databases (GSDs) are submitted in a wide variety of data formats to CoL by hundreds of taxonomic experts and institutions. Submitted data are reformatted to a standard data submission format: CoL Standard Dataset (ACEF), DarwinCore, or CoL DP. A series of standardized data integrity checks are run to detect and resolve frequently occurring data quality problems including character encoding corruption, non-Latin characters in scientific names, missing parents, duplicated and homonymic names within the GSD and among other GSDs, split taxonomic groups that have been assigned to multiple parent taxa, and other issues. The process and challenges of assembling data into the Catalogue of Life, and future directions of the project in migrating to CoL+ infrastructure will be discussed.

© Ower G, Roskov Y. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Keywords

Catalogue of Life, global taxonomic checklist, biodiversity informatics, taxonomy, nomenclature, species richness

Presenting author

Geoffrey Ower

Presented at

Biodiversity_Next 2019